On progressively censored generalized exponential distribution

Biswabrata Pradhan · Debasis Kundu

Abstract In this paper, we consider the statistical inference of the unknown parameters of the generalized exponential distribution in presence of progressive censoring. We obtain maximum likelihood estimators of the unknown parameters using EM algorithm. We also compute the expected Fisher information matrix using the missing value principle. We then use these values to determine the optimal progressive censoring plans. Different optimality criteria are considered, and selected optimal progressive censoring plans are presented. One example has been provided for illustrative purposes.

Keywords Maximum likelihood estimators · EM algorithm · Fisher information matrix · Asymptotic distribution · Optimal censoring scheme

Mathematics Subject Classification (2000) 62N05 · 62F10

1 Introduction

Recently, the Type-II progressively censoring scheme has received considerable interest among the statisticians. It can be described as follows. Suppose that n units are placed on a life test and the experimenter decides beforehand the quantity m, the number of units to be failed. Now at the time of the first failure, R_1 of the remaining $n-1$ surviving units are randomly removed from the experiment. Continuing

B. Pradhan (✉)
SQC & OR Unit, Indian Statistical Institute, 203 B.T. Road, Kolkata, Pin 700108, India
e-mail: bis@isical.ac.in

D. Kundu
Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India
e-mail: kundu@iitk.ac.in

on, at the time of the second failure, R_2 of the remaining $n - R_1 - 2$ units are randomly withdrawn from the experiment. Finally, at the time of the mth failure, all the remaining $n - m - R_1 - \cdots - R_{m-1}(= R_m)$ surviving units are removed from the experiment. Some of the earlier work on progressive censoring was conducted by Cohen (1963), Mann (1971), and Thomas and Wilson (1972). Recently, several articles have been published on estimating the parameters of the unknown parameters for different distribution functions, see, for example, Viveros and Balakrishnan (1994), Balasooriya and Balakrishnan (2000), Balakrishnan et al. (2004), Balakrishnan and Kannan (2001), etc. A recent account on progressive censoring schemes can be obtained in the monograph by Balakrishnan and Aggarwala (2000) or in the excellent review article by Balakrishnan (2007).

Although quite a bit of work has been done on the progressively censored Weibull distribution, we have not come across any work on the progressively censored gamma or generalized exponential (GE) distribution. It is observed (Gupta and Kundu 1999) that the GE distribution can be used quite effectively to analyze lifetime data in place of the Weibull or gamma distribution. The two-parameter GE distribution has the density/hazard functions which are very similar to the density/hazard functions of the two-parameter gamma distribution, see Gupta and Kundu (1999). Since the distribution function of the GE distribution is also in a compact form like the Weibull distribution, it is observed that it can be used very easily when the data are censored, unlike the gamma distribution.

In this paper, we consider the statistical inference of the GE distribution when the data are progressively censored. We obtain likelihood equations, and it is observed that the maximum likelihood estimators (MLEs) can not be obtained in explicit forms. The MLEs can be obtained by solving a two-dimensional optimization problem. It is observed that in certain cases the standard Newton–Raphson algorithm does not converge. We propose to use the EM algorithm to compute the MLEs. We also calculate the expected Fisher information matrix using the missing information principle, and they have been used for constructing asymptotic confidence intervals.

Finding any optimal sampling scheme is an important practical problem and it has received considerable attention in the last few years. But most of the work is related to the progressively censored Weibull distribution. Finding an optimal censoring plan means the choice of $(R_1, \ldots, R_m)$ for specified values of sample size n and effective sample size m which provides the maximum information. In this paper, we consider an optimality criterion (measure of information) based on the expected Fisher information matrix and, using this criterion, we propose a method to choose an optimal sampling scheme for progressively censored GE distribution. Monte Carlo simulations are performed to study the behavior of the proposed methodology, and two data sets are analyzed for illustrative purposes.

The rest of the paper is organized as follows. In Sect. 2, we provide the maximum likelihood estimators using the EM algorithm. The observed and expected Fisher information matrices are provided in Sect. 3. Numerical simulations and data analysis are provided in Sect. 4. In Sect. 5, we provide a procedure to obtain the optimum censoring scheme and, finally, we conclude the paper in Sect. 6.

2 Maximum likelihood estimators

2.1 Model description

A random variable X is said to have two-parameter GE distribution if the probability density function (PDF) is of the following form:

$$f(x|\alpha, \lambda) = \begin{cases} 0 & \text{if } x < 0, \\ \alpha\lambda e^{-\lambda x}(1 - e^{-\lambda x})^{\alpha - 1} & \text{if } x \geq 0 \end{cases} \tag{1}$$

for $\alpha > 0$ and $\lambda > 0$. Here α and λ are the shape and scale parameters, respectively. It is known that, for $\alpha \leq 1$, the PDF is a decreasing function and, for $\alpha > 1$, it is a unimodal function. Shapes of the different PDF of GE distributions can be found in Gupta and Kundu (1999). For some recent developments of the GE distribution, the readers may look at the review article by Gupta and Kundu (2007).

2.2 Maximum likelihood estimators

Suppose that n independent items are put on a test and that the lifetime distribution of each item is given by (1). The ordered m-failures are observed under the type-II progressively censoring plan $(R_1, \ldots, R_m)$, where each $R_i \geq 0$ and $\sum_{j=1}^{m} R_j + m = n$. If the ordered m failures are denoted by $x_{(1)} < \cdots < x_{(m)}$, then the likelihood function based on the observed sample $x_{(1)} < \cdots < x_{(m)}$ is

$$l(\alpha, \lambda) = c \prod_{i=1}^{m} f(x_{(i)}|\alpha, \lambda)\big[1 - F(x_{(i)}|\alpha, \lambda)\big]^{R_i}, \tag{2}$$

where $c = n(n - 1 - R_1) \cdots (n - R_1 - \cdots - R_{m-1} - m + 1)$ and $F(x|\alpha, \lambda) = (1 - e^{-\lambda x})^{\alpha}$, the distribution function corresponding to the density function (1).

The MLEs of α and λ can be obtained by solving two nonlinear normal equations, whose explicit solutions cannot be obtained. They have to be obtained by solving a two-dimensional optimization problem. It is observed that the standard Newton–Raphson algorithm does not converge in certain cases. We propose to use the EM algorithm to compute the MLEs of α and λ as suggested by Dempster et al. (1977) which involves solving two one-dimensional optimization problems rather than one two-dimensional problem.

The progressive right censoring model problem can be viewed as an incomplete data problem, see, for example, Ng et al. (2002). First, let us denote the observed and censored data by $X = (X_{(1)}, \ldots, X_{(m)})$ and $Z = (Z_1, \ldots, Z_m)$, respectively, where each Z_j is $1 \times R_j$ vector with $Z_j = (Z_{j1}, \ldots, Z_{jR_j})$ for $j = 1, \ldots, m$, and they are not observable. The censored data vector Z can be thought of as missing data. The combination of $(X, Z) = W$ forms the complete data set. If we denote the log-likelihood function of the uncensored data set by $L_c(W; \alpha, \lambda)$, then, ignoring the additive constant, we have

$$L_c(W; \alpha, \lambda) = n \ln \alpha + n \ln \lambda - \lambda \sum_{i=1}^{m} x_{(i)} + (\alpha - 1) \sum_{i=1}^{m} \ln\big(1 - e^{-\lambda x_{(i)}}\big)$$

$$- \lambda \sum_{j=1}^{m} \sum_{k=1}^{R_j} z_{jk} + (\alpha - 1) \sum_{j=1}^{m} \sum_{k=1}^{R_j} \ln\left(1 - e^{-\lambda z_{jk}}\right). \tag{3}$$

For the E-step, one needs to compute the pseudo log-likelihood function. It can be obtained from $L_c(W; \alpha, \lambda)$ by replacing any function of z_{jk}, say $g(z_{jk})$, with $E\{g(Z_{jk})|Z_{jk} > x_{(j)}\}$. Therefore, the pseudo log-likelihood function becomes

$$n \ln \alpha + n \ln \lambda - \lambda \sum_{i=1}^{m} x_{(i)} + (\alpha - 1) \sum_{i=1}^{m} \ln\left(1 - e^{-\lambda x_{(i)}}\right)$$

$$- \lambda \sum_{j=1}^{m} \sum_{k=1}^{R_j} E(Z_{jk}|Z_{jk} > x_{(j)})$$

$$+ (\alpha - 1) \sum_{j=1}^{m} \sum_{k=1}^{R_j} E\left[\ln\left(1 - e^{-\lambda Z_{jk}}\right)|Z_{jk} > x_{(j)}\right]. \tag{4}$$

It can be seen that (see the Appendix)

$$A(c, \alpha, \lambda) = E(Z_{jk}|Z_{jk} > c) = -\frac{\alpha}{\lambda(1 - F(a|\alpha, \lambda))} \times u(c, \alpha, \lambda), \tag{5}$$

where

$$u(c, \alpha, \lambda) = \int_{0}^{e^{-c\lambda}} \ln y (1 - y)^{\alpha - 1} \, dy$$

and

$$B(c, \alpha, \lambda) = E\left[\ln\left(1 - e^{-\lambda Z_{jk}}\right)\Big|Z_{jk} > c\right]$$

$$= \frac{1}{\alpha(1 - (1 - e^{-c\lambda})^{\alpha})} \times \left\{\left(1 - e^{-c\lambda}\right)^{\alpha}\left(1 - \ln\left(1 - e^{-c\lambda}\right)^{\alpha}\right) - 1\right\}. \tag{6}$$

Now the M-step involves the maximization of the pseudo log-likelihood function (4), replacing the corresponding values of (5) and (6) in (4). Therefore, if at the kth stage, the estimate of (α, λ) is $(\alpha^{(k)}, \lambda^{(k)})$, then $(\alpha^{(k+1)}, \lambda^{(k+1)})$ can be obtained by maximizing

$$L_c^*(W; \alpha, \lambda) = n \ln \alpha + n \ln \lambda - \lambda \sum_{i=1}^{m} x_{(i)} + (\alpha - 1) \sum_{i=1}^{m} \ln\left(1 - e^{-\lambda x_{(i)}}\right)$$

$$- \lambda \sum_{j=1}^{m} R_j A\left(x_{(j)}, \alpha^{(k)}, \lambda^{(k)}\right)$$

$$+ (\alpha - 1) \sum_{j=1}^{m} R_j B\left(x_{(j)}, \alpha^{(k)}, \lambda^{(k)}\right) \tag{7}$$

with respect to α and λ. Note that the maximization of (7) can be obtained quite effectively by the similar method proposed by Gupta and Kundu (2001). First, find $\lambda^{(k+1)}$ by solving a fixed-point type equation

$$h(\lambda) = \lambda, \tag{8}$$

where the function $h(\lambda)$ is defined by

$$h(\lambda) = \left[\frac{1}{n} \sum_{j=1}^{m} x_{(j)} + \frac{1}{n} \widetilde{A} - \frac{1}{n} \big(\widehat{\alpha}(\lambda) - 1\big) \sum_{j=1}^{m} \frac{x_{(j)} e^{-\lambda x_{(j)}}}{1 - e^{-\lambda x_{(j)}}} \right]^{-1}$$

with

$$\widetilde{A} = \sum_{j=1}^{m} R_j A\big(x_{(j)}, \alpha^{(k)}, \lambda^{(k)}\big), \qquad \widetilde{B} = \sum_{j=1}^{m} R_j B\big(x_{(j)}, \alpha^{(k)}, \lambda^{(k)}\big),$$

$$\widehat{\alpha}(\lambda) = -\frac{n}{\sum_{j=1}^{m} \ln(1 - e^{-\lambda x_{(j)}}) + \widetilde{B}}.$$

Once $\lambda^{(k+1)}$ is obtained, $\alpha^{(k+1)}$ is obtained as $\alpha^{(k+1)} = \widehat{\alpha}(\lambda^{(k+1)})$. Therefore, we can use the following algorithm to proceed from the kth iterate to $(k + 1)$th iterate.
Algorithm

Step 1: Maximize (7) using (8), i.e., continue the process until it converges. At the $(k + 1)$th stage, the value of λ that maximizes (7) is $\lambda^{(k+1)}$.
Step 2: Assign

$$\alpha^{(k+1)} = -\frac{n}{\sum_{j=1}^{m} \ln(1 - e^{-\lambda^{(k+1)} x_{(j)}}) + \widetilde{B}}.$$

Step 3: Check the convergence of $(\alpha^{(k+1)}, \lambda^{(k+1)})$ If the convergence is met, stop the iteration, otherwise go back to Step 1.

3 Fisher information matrix

In this section, we compute the observed and expected Fisher information matrices using the idea of missing information principle of Louis (1982); see also Tanner (1993). The observed Fisher information matrix can be used to construct the asymptotic confidence intervals, whereas the expected Fisher information matrix will be used for constructing optimal censoring plans.

3.1 Observed information matrix

In this subsection, we compute the observed Fisher information matrix, given the observations $\{x_{(1)}, \ldots, x_{(m)}\}$ and $R_1, \ldots, R_m$. The idea of the missing information principle of Louis (1982) can be expressed as follows:

Observed information = Complete information − Missing information.

In our case, if we denote $\theta = (\alpha, \lambda)$, X = the observed data, W = the complete data, $I_W(\theta)$ = the complete information, $I_X(\theta)$ = the observed information, and $I_{W|X}(\theta)$ = the missing information, then they can be expressed as follows:

$$I_X(\theta) = I_W(\theta) - I_{W|X}(\theta). \tag{9}$$

The complete information $I_W(\theta)$ is given by

$$I_W(\theta) = -E\left[\frac{\partial^2 L_c(W; \theta)}{\partial \theta^2}\right].$$

The Fisher information in one observation, which is censored at the time of the jth failure time $x_{(j)}$, can be computed as

$$I_{W|X}^{(j)}(\theta) = -E_{Z_j|X_{(j)}}\left[\frac{\partial^2 \ln f_{Z_j}(z_j|x_{(j)}, \theta)}{\partial \theta^2}\right].$$

Therefore, the expected information for conditional distribution of W given X (the missing information) is

$$I_{W|X}(\theta) = \sum_{j=1}^{m} R_j I_{W|X}^{(j)}(\theta).$$

So the observed information can be obtained from (9). The asymptotic variance covariance matrix of $\widehat{\theta} = (\widehat{\alpha}, \widehat{\lambda})$ can be obtained by inverting $I_X(\widehat{\theta})$.

Both the matrices $I_W(\theta)$ and $I_{W|X}(\theta)$ are of order 2×2. Now, we present all the elements of both matrices. The 2×2 matrix $I_W(\theta)$ for complete data is already available, see Gupta and Kundu (2001). For convenience, we present it below. If we denote the (i, j)th element of the matrix $I_W(\theta)$ by $a_{ij}(\alpha, \lambda)$, then they are as follows:

$$a_{11} = \frac{n}{\alpha^2},$$

$$a_{12} = a_{21} = -\frac{n}{\lambda}\left[\frac{\alpha}{\alpha - 1}(\psi(\alpha) - \psi(1)) - (\psi(\alpha + 1) - \psi(1))\right] \quad \text{if } \alpha > 2,$$

$$= -\frac{n\alpha}{\lambda}\int_0^\infty xe^{-2x}(1 - e^{-x})^{\alpha - 2}dx \quad \text{if } 0 < \alpha \le 2,$$

$$a_{22} = \frac{n}{\lambda^2}\left[1 + \frac{\alpha(\alpha - 1)}{\alpha - 2}\left(\psi'(1) - \psi'(\alpha - 1) + (\psi(\alpha - 1) - \psi(1))^2\right)\right]$$

$$\quad + \frac{n\alpha}{\lambda^2}\left[(\psi'(1) - \psi(\alpha)) + (\psi(\alpha) - \psi(1))^2\right] \quad \text{if } \alpha > 2,$$

$$= \frac{n}{\lambda^2} + \frac{n\alpha(\alpha - 1)}{\lambda^2}\int_0^\infty x^2 e^{-2x}(1 - e^{-x})^{\alpha - 3}dx \quad \text{if } 0 < \alpha \le 2.$$

Here ψ and ψ' are the digamma and trigamma functions, see, for example, Abramowitz and Stegun (1964).

Now we present $I_{W|X}^{(j)}(\theta)$. If

$$I_{W|X}^{(j)}(\theta) = \begin{bmatrix} b_{11}(x_{(j)}; \alpha, \lambda) & b_{12}(x_{(j)}; \alpha, \lambda) \\ b_{21}(x_{(j)}; \alpha, \lambda) & b_{22}(x_{(j)}; \alpha, \lambda) \end{bmatrix},$$

then

$$b_{11}(x_{(j)}; \alpha, \lambda) = \frac{1}{\alpha^2} - \left[\ln\left(1 - e^{-\lambda x_{(j)}}\right)\right]^2 \frac{(1 - e^{-\lambda x_{(j)}})^\alpha}{(1 - (1 - e^{-\lambda x_{(j)}})^\alpha)^2},$$

$$b_{22}(x_{(j)}; \alpha, \lambda) = \frac{1}{\lambda^2} + (\alpha - 1)h_1(x_{(j)}; \alpha, \lambda)$$

$$- \frac{\alpha x_{(j)}^2 e^{-\lambda x_{(j)}}(1 - e^{-\lambda x_{(j)}})^{\alpha-2}}{(1 - (1 - e^{-\lambda x_{(j)}})^\alpha)^2} \times \left[\alpha e^{-\lambda x_{(j)}} - 1 + \left(1 - e^{-\lambda x_{(j)}}\right)^\alpha\right]$$

$$b_{12}(x_{(j)}; \alpha, \lambda) = -h_2(x_{(j)}; \alpha, \lambda) + \frac{x_{(j)} e^{-\lambda x_{(j)}}(1 - e^{-\lambda x_{(j)}})^{\alpha-1}}{(1 - (1 - e^{-\lambda x_{(j)}})^\alpha)^2}$$

$$\times \left[1 + \alpha \ln\left(1 - e^{-\lambda x_{(j)}}\right) - \left(1 - e^{-\lambda x_{(j)}}\right)^\alpha\right],$$

where

$$h_1(x_{(j)}; \alpha, \lambda) = \frac{1}{\lambda^2(1 - (1 - e^{-\lambda x_{(j)}})^\alpha)}$$

$$\times \int_{(1-e^{-\lambda x_{(j)}})^\alpha}^1 \left(\ln\left(1 - u^{1/\alpha}\right)\right)^2 \left(1 - u^{1/\alpha}\right)u^{-2/\alpha}\, du \quad \text{and}$$

$$h_2(x_{(j)}; \alpha, \lambda) = \frac{1}{\lambda(1 - (1 - e^{-\lambda x_{(j)}})^\alpha)}$$

$$\times \int_{(1-e^{-\lambda x_{(j)}})^\alpha}^1 \left(-\ln\left(1 - u^{1/\alpha}\right)\right)\left(1 - u^{1/\alpha}\right)u^{-1/\alpha}\, du.$$

3.2 Expected Fisher information matrix

In this subsection, we provide the expected Fisher information matrix for the progressively censored data. For this, we need the following. The probability density function of $X_{(j)}$ for $j = 1, \ldots, m$ is

$$f_{X_{(j)}}(x) = c_{j-1} \sum_{i=1}^j a_{i,j} \left(1 - \left(1 - e^{-\lambda x}\right)^\alpha\right)^{r_i - 1} \alpha \lambda e^{-\lambda x}\left(1 - e^{-\lambda x}\right)^{\alpha-1} \tag{10}$$

for $x > 0$ and 0 otherwise, where

$$r_j = m - j + 1 + \sum_{i=j}^m R_i, \qquad c_{j-1} = \prod_{i=1}^j r_i \quad \text{for } j = 1, \ldots, m$$

and

$$a_{11} = 1, \qquad a_{i,j} = \prod_{k=1, k \neq i}^{j} \frac{1}{r_k - r_i} \quad \text{for } 1 \leq i \leq j \leq m.$$

See, for example, Balakrishnan and Aggarwala (2000), p. 26.

Based on (10), the expected Fisher information matrix can be obtained. If we denote the 2×2 matrix E by

$$E = \begin{bmatrix} E_{11} & E_{12} \\ E_{21} & E_{22} \end{bmatrix} = -E \begin{bmatrix} \frac{\partial^2 \ln l(\alpha, \lambda)}{\partial \alpha^2} & \frac{\partial^2 \ln l(\alpha, \lambda)}{\partial \alpha \partial \lambda} \\ \frac{\partial^2 \ln l(\alpha, \lambda)}{\partial \lambda \partial \alpha} & \frac{\partial^2 \ln l(\alpha, \lambda)}{\partial \lambda^2} \end{bmatrix},$$

then

$$E_{11} = \frac{1}{\alpha^2} \left[m + \sum_{j=1}^{m} R_j h_{1j}(\alpha, \lambda) \right],$$

$$E_{22} = \frac{1}{\lambda^2} \left[m + \alpha(\alpha - 1) \sum_{j=1}^{m} h_{2j}(\alpha, \lambda) + \alpha^2 \sum_{j=1}^{m} h_{3j}(\alpha, \lambda) \right],$$

$$E_{12} = E_{21} = -\frac{\alpha}{\lambda} \left[\sum_{j=1}^{m} h_{4j}(\alpha, \lambda) - \sum_{j=1}^{m} R_j h_{5j}(\alpha, \lambda) - \alpha \sum_{j=1}^{m} R_j h_{6j}(\alpha, \lambda) \right],$$

where

$$h_{1j}(\alpha, \lambda) = c_{j-1} \sum_{i=1}^{j} a_{i,j} \int_0^1 u(\ln u)^2 (1-u)^{r_i - 3} \, du,$$

$$h_{2j}(\alpha, \lambda) = c_{j-1} \sum_{i=1}^{j} a_{i,j} \int_0^1 u^{\alpha - 3} (1 - u^\alpha)^{r_i - 1} (1-u) (\ln(1-u))^2 \, du,$$

$$h_{3j}(\alpha, \lambda) = c_{j-1} \sum_{i=1}^{j} a_{i,j} \int_0^1 u^{2\alpha - 3} (1 - u^\alpha)^{r_i - 3} (1-u) (\ln(1-u))^2$$
$$\times (\alpha u - u^\alpha - \alpha + 1) \, du,$$

$$h_{4j}(\alpha, \lambda) = c_{j-1} \sum_{i=1}^{j} a_{i,j} \int_0^1 u^{\alpha - 2} (1 - u^\alpha)^{r_i - 3} (1-u) (-\ln(1-u)) \, du,$$

$$h_{5j}(\alpha, \lambda) = c_{j-1} \sum_{i=1}^{j} a_{i,j} \int_0^1 u^{2\alpha - 2} (1 - u^\alpha)^{r_i - 2} (1-u) (-\ln(1-u)) \, du,$$

$$h_{5j}(\alpha, \lambda) = c_{j-1} \sum_{i=1}^{j} a_{i,j} \int_0^1 u^{2\alpha - 2} (1 - u^\alpha)^{r_i - 3} (1-u) (-\ln(1-u)) \ln u \, du.$$

Therefore, the asymptotic variance covariance matrix of the MLEs of (α, λ) becomes

$$V = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} = E^{-1}.$$

4 Numerical experiments and data analysis

4.1 Experimental results

In this subsection, we present some experimental results to observe how the MLEs perform for different sampling schemes and for different sample sizes. All the computations are performed at the Indian Statistical Institute Kolkata. We have taken $n = 20, 25, 30$, $m = 10$ and 15, and eleven ([1]–[11]) different sampling schemes. In all cases, we have used $\lambda = 1.0$ and we have taken $\alpha = 1.5$ and $\alpha = 0.75$. For particular n, m, and a sampling scheme, we have used the method proposed by Balakrishnan and Sandhu (1995) to generate progressively censored generalized exponential samples. In each case, we have calculated the MLEs using the EM algorithm and also the asymptotic confidence intervals of α and λ based on the Fisher information matrix. We replicate the process 10000 times and compute the average biases and standard deviations of the different estimates.

Note that the sampling schemes [1], [3], [6], [8], [10] are the usual Type-II censoring scheme, i.e., $n - m$ items are removed at the time of the mth failure. The sampling schemes [2], [4], [7], [9], [11] are just the opposite of the Type-II sampling schemes, i.e., $n - m$ items are removed at the time of the first failure. In this paper, we will refer this censoring scheme as Type-III censoring scheme. It is well known that, for fixed n and m, the expected experimental time of the Type-II censoring schemes is less than that of the corresponding Type-III schemes. In fact, the expected time of any other censoring scheme, for fixed n and m, will be always between these two extremes. For illustrative purpose, we have taken one arbitrary censoring scheme, scheme [5], whose expected experimental time is between the schemes [3] and [4].

For different sampling schemes, we use the notation of Ng et al. (2004). For example, when $n = 25$ and $m = 10$, the scheme (5, 5, 5, 7*0) means that, after the first failure, from the remaining 24 items 5 items are removed at random, after the second failure, from the remaining 18 items 5 items are removed at random, similarly, after the third failure, from the remaining 12 items 5 items are removed at random, and we observe the next seven failure times. The results are reported in Tables 1–4.

Some of the points are quite clear from Tables 1 and 2. For example, in all the cases considered, the average biases and the standard deviations of the MLEs of α and λ are smallest for Type-III schemes and largest for Type-II schemes. For other censoring schemes, the average biases and the standard deviations are between these two extremes. It is also observed that, for fixed m, as n increases, the average biases and the standard deviations decrease. Moreover, for fixed n, as m increases, the same phenomena are observed. Therefore, it is clear that, as the sample size increases or the effective sample size increases, the performances of the MLEs in terms of biases and standard deviations become better. When the effective sample size is 15, then, in most

Table 1 The average biases (AB), the standard deviations (SD), 95% coverage percentages (CP-95), and 90% coverage percentages (CP-90) for MLE of α when $\alpha = 1.5$ are presented for different sample sizes and different sampling schemes

n	m	Scheme	No.	AB	SD	CP-95	CP-90
20	10	(9*0, 10)	[1]	−0.1537	0.3024	92.3	88.2
20	10	(10, 9*0)	[2]	−0.1382	0.2737	95.2	89.7
25	10	(9*0, 15)	[3]	−0.2599	0.2832	93.8	88.7
25	10	(15, 9*0)	[4]	−0.2303	0.2610	94.1	90.5
25	10	(5, 5, 5, 7*0)	[5]	−0.2493	0.2730	94.2	91.8
25	15	(14*0, 10)	[6]	−0.0980	0.2156	96.8	91.2
25	15	(10, 14*0)	[7]	−0.0902	0.2025	97.1	92.5
30	10	(9*0, 20)	[8]	−0.3384	0.2730	93.5	87.2
30	10	(20, 9*0)	[9]	−0.2840	0.2388	94.2	87.8
30	15	(14*0, 15)	[10]	−0.1880	0.2309	97.0	92.5
30	15	(15, 14*0)	[11]	−0.1760	0.2127	96.2	91.5

Table 2 The average biases (AB), the standard deviations (SD), 95% coverage percentages (CP-95), and 90% coverage percentages (CP-90) for MLE of λ when $\alpha = 1.5$ are presented for different sample sizes and different sampling schemes

n	m	Scheme	No.	AB	SD	CP-95	CP-90
20	10	(9*0, 10)	[1]	0.1509	0.2786	91.2	85.6
20	10	(10, 9*0)	[2]	−0.1400	0.2537	92.3	87.8
25	10	(9*0, 15)	[3]	−0.2410	0.2240	93.8	88.7
25	10	(15, 9*0)	[4]	−0.2128	0.2128	92.5	89.2
25	10	(5, 5, 5, 7*0)	[5]	−0.2202	0.2172	93.6	89.5
25	15	(14*0, 10)	[6]	−0.0997	0.2693	92.7	88.1
25	15	(10, 14*0)	[7]	−0.0342	0.2156	94.2	89.6
30	10	(9*0, 20)	[8]	−0.3150	0.2130	92.5	86.5
30	10	(20, 9*0)	[9]	−0.2649	0.1799	93.2	87.8
30	15	(14*0, 15)	[10]	−0.1486	0.2072	95.0	91.0
30	15	(15, 14*0)	[11]	−0.1054	0.1510	96.2	90.5

of the cases, the coverage percentages are very close to the corresponding nominal levels. Therefore, the asymptotic results can be used for all practical purposes.

One of the referees raised a valid point that why an EM algorithm should be used rather than the traditional Newton–Raphson method. It may be mentioned that to employ the Newton–Raphson method, one needs to compute the second derivatives of the log-likelihood function. Sometimes, like in the present situation, calculation of the second derivatives based on progressive censored data are quite complicated. In such a scenario, the EM algorithm is a very useful technique. Moreover, as Little and Rubin (1983) correctly mentioned, the EM algorithm will converge rather slowly but more reliably (as compared to the Newton–Raphson method) when the amount of information in the missing data is relatively large. We present a small comparison be-

Table 3 The average biases (AB), the standard deviations (SD), 95% coverage percentages (CP-95), and 90% coverage percentages (CP-90) for MLE of α when $\alpha = 0.75$ are presented for different sample sizes and different sampling schemes

n	m	Scheme	No.	AB	SD	CP-95	CP-90
20	10	(9*0, 10)	[1]	0.2090	0.2445	91.6	88.9
20	10	(10, 9*0)	[2]	0.1467	0.2207	94.9	90.2
25	10	(9*0, 15)	[3]	0.2036	0.1853	92.8	85.8
25	10	(15, 9*0)	[4]	0.1421	0.1719	92.6	88.2
25	10	(5, 5, 5, 7*0)	[5]	0.1470	0.1792	94.5	90.5
25	15	(14*0, 10)	[6]	0.1334	0.2342	93.2	89.7
25	15	(10, 14*0)	[7]	0.1271	0.2107	96.2	92.5
30	10	(9*0, 20)	[8]	0.2058	0.1634	92.2	87.6
30	10	(20, 9*0)	[9]	0.1390	0.1316	92.8	89.2
30	15	(14*0, 15)	[10]	0.1627	0.1449	95.8	89.8
30	15	(15, 14*0)	[11]	0.1208	0.1126	96.8	91.3

Table 4 The average biases (AB), the standard deviations (SD), 95% coverage percentages (CP-95), and 90% coverage percentages (CP-90) for MLE of λ when $\alpha = 0.75$ are presented for different sample sizes and different sampling schemes

n	m	Scheme	No.	AB	SD	CP-95	CP-90
20	10	(9*0, 10)	[1]	0.6104	0.7852	90.5	84.9
20	10	(10, 9*0)	[2]	0.4448	0.6281	91.8	85.2
25	10	(9*0, 15)	[3]	0.7190	0.7551	88.1	84.9
25	10	(15, 9*0)	[4]	0.5250	0.6128	88.8	83.2
25	10	(5, 5, 5, 7*0)	[5]	0.5541	0.6888	90.2	86.8
25	15	(14*0, 10)	[6]	0.3931	0.5426	90.2	86.2
25	15	(10, 14*0)	[7]	0.3138	0.4379	92.8	88.0
30	10	(9*0, 20)	[8]	0.8248	0.7312	91.2	86.6
30	10	(20, 9*0)	[9]	0.5886	0.6013	92.5	88.5
30	15	(14*0, 15)	[10]	0.4795	0.5498	93.7	89.1
30	15	(15, 14*0)	[11]	0.3680	0.4482	94.5	90.5

tween the EM algorithm and the Newton–Raphson method in the following Table 5. In Table 5, we present the number of times (out of 1000) the Newton–Raphson algorithm converges (NTC). In all the cases, the EM algorithm has always converged. It clearly shows the advantage of using the EM algorithm.

4.2 Data analysis

In this subsection, we provide a data analysis for illustrative purposes. The data have been taken from Lawless (1982, p. 491), and it represents the failure or censoring times of 36 appliances subjected to an automatic life test. The data given below consist of only the failure times: 11, 35, 49, 170, 329, 381, 708, 958, 1062, 1167, 1594,

Table 5 The number of times (NTC) the Newton–Raphson algorithm converges out of 1000 replications for different censoring schemes and different sets of parameters

n	m	Scheme	α	λ	NTC
20	10	(10, 9*0)	1.5	1.0	921
20	10	(10, 9*0)	0.75	1.0	944
25	15	(10, 14*0)	1.5	1.0	957
25	15	(10, 14*0)	0.75	1.0	971

1925, 1990, 2223, 2327, 2400, 2451, 2471, 2551, 2565, 2568, 2694, 2702, 2761, 2831, 3034, 3059, 3112, 3214, 3478, 3504, 4329, 6367, 6976, 7846, 13403.

Before progressing further, we have first fitted the GE distribution to the complete data set, and it is observed that $\widehat{\alpha} = 0.96001$ and $\widehat{\lambda} = 0.00035$. The Kolmogorov–Smirnov distance is 0.202, and the corresponding p value is 0.11. Therefore, the GE distribution provides a reasonable fit. Moreover, we have tested the following hypothesis: H_0 : Data follow exponential, $vs.$ H_1 : Data follow GE. The $-2\ln(L_0 - L_1) = 7.013$, and it implies that H_0 is rejected with the level of significance less than 0.001.

We have generated progressively censored samples using three different sampling schemes, from the above data with $m = 12$, as follows:

Censoring Scheme 1: (15, 5, 4, 9*0). We obtain the following progressively censored sample: 11, 35, 49, 329, 1062, 1167, 1594, 1990, 2451, 2471, 2551, 3059.

Censoring Scheme 2: (11*0, 24). We obtain the following progressively censored sample: 11, 35, 49, 170, 329, 381, 708, 958, 1062, 1167, 1594, 1925.

Censoring Scheme 3: (24, 11*0). We obtain the following progressively censored sample: 11, 35, 49, 329, 381, 958, 1062, 1594, 1925, 2223, 2451, 2471.

In all the three cases, we have used the EM algorithm to compute the MLEs. The MLEs of (α, λ) for Scheme-1, Scheme-2, and Scheme-3 are (0.89532, 0.00074), (0.79080, 0.00020), and (0.88723, 0.00093), respectively. The corresponding variance-covariance matrices are

$$\begin{bmatrix} 6.217 \times 10^{-2} & 4.143 \times 10^{-5} \\ 4.143 \times 10^{-5} & 7.767 \times 10^{-8} \end{bmatrix}, \quad \begin{bmatrix} 9.338 \times 10^{-2} & 4.356 \times 10^{-5} \\ 4.356 \times 10^{-5} & 2.522 \times 10^{-8} \end{bmatrix},$$

and

$$\begin{bmatrix} 6.313 \times 10^{-2} & 5.192 \times 10^{-5} \\ 5.192 \times 10^{-5} & 1.237 \times 10^{-7} \end{bmatrix},$$

respectively. In Fig. 1, we have plotted the estimated distribution functions based on complete sample and also based on three different sampling schemes. Interestingly, in this case, the estimated distribution function based on Scheme-2 (Type-II censoring) provides the closest approximation to the distribution function based on complete sample.
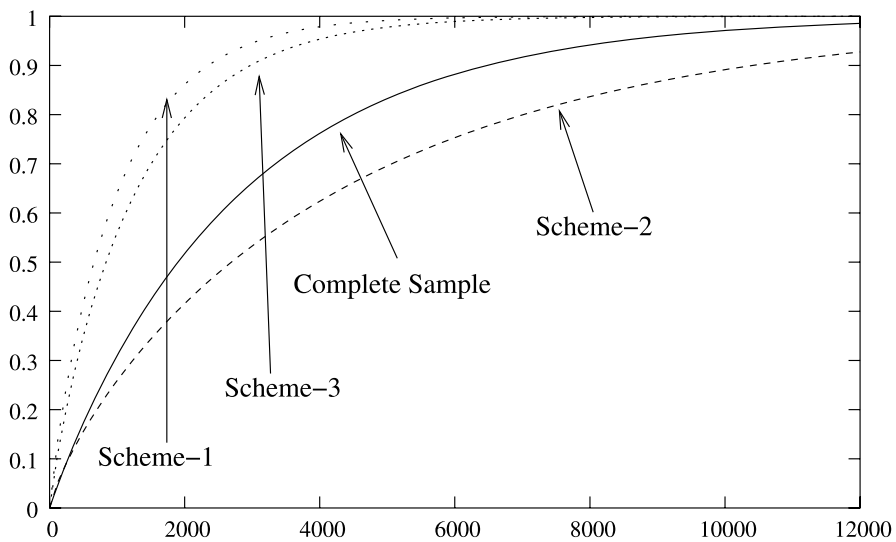
Fig. 1 Four different estimated distribution function

5 Optimal censoring scheme

So far we have discussed about the statistical inferences of the unknown parameters of the GE distributions when the data are progressively censored for a known censoring scheme. A natural question arises how to choose a particular censoring scheme. Should we choose a particular scheme just based on convenience or based on some statistical criteria. Recently, choosing the optimal censoring scheme in different problems has received considerable attention in the statistical literature. See, for example, Zhang and Meeker (2005), Ng et al. (2004), Kundu (2007), Wang and Yu (2007), and the references cited there.

For a practitioner, it is quite important to choose the *optimal* censoring scheme from a class of possible schemes. Here possible schemes mean, for fixed sample size n and for fixed effective sample size m, the different choices of $R_1, \ldots, R_m$, such that

$$R_1 + \cdots + R_m + m = n. \tag{11}$$

In many practical situations, the experimenter may not have any choice on m and n, but he/she can choose a particular $(R_1, \ldots, R_m)$ satisfying (11). Therefore, our problem boils down as follows: for fixed m and n, choose that particular scheme $(R_1, \ldots, R_m)$ which is *optimal* in the sense it provides the maximum information of the unknown parameters.

Immediately, the first question arises how to define the information measure of the unknown parameters for a particular censoring scheme, or how to compare two censoring schemes based on their information measures? In this respect, comparing the Fisher information matrices seems to be a natural choice. If the model has only one unknown parameter, then this comparison can be easily made. But, if both the parameters are unknown, then the comparison of the two Fisher information matrices

is not a trivial task. Some of the existing choices are to compare the traces or the determinants of the two Fisher information matrices. Unfortunately, in presence of the shape and scale parameters, it can be easily seen that (Gupta and Kundu 2006) the trace or the determinant is not scale invariant. Therefore, it may happen that, for a particular scheme, its determinant or trace of the Fisher information matrix is more than another scheme, but if we change the unit of the data (multiply the data by a positive constant), then the inequality becomes reversed, which may not be very desirable.

An alternative way of comparing the information measures of two different schemes is to compare their precisions of the $100p$th quantile estimators, i.e., to compare the variances of the corresponding estimators for different schemes. Similar ideas were used by Zhang and Meeker (2005) in the Bayesian set up and also by Ng et al. (2004) in the frequentest context. Interestingly, this information measure is independent of the scale parameter, but unfortunately it depends on 'p'. Balakrishnan and Aggarwala (2000) proposed for some specific choices of 'p' based on some practical consideration, may be $p = 0.95$ or $p = 0.99$, but they can be argued upon.

Here we use the following information measure. Consider the pth quantile of the lifetime distribution,

$$T_p = -\frac{1}{\lambda} \ln\left(1 - p^{\frac{1}{\alpha}}\right).$$

In this paper, following the idea of Gupta and Kundu (2006), we consider the following information measures for a particular censoring scheme:

$$I_W\left\{(R_1, \ldots, R_m)\right\} = \int_0^1 V\left\{(R_1, \ldots, R_m)\right\}_p W(p)\, dp, \tag{12}$$

where $V\{(R_1, \ldots, R_m)\}_p$ denotes the asymptotic variance of $\widehat{T_p}$, the MLE of T_p, based on the censoring scheme $(R_1, \ldots, R_m)$. Here $W(p) \geq 0$ is a nonnegative weight function such that

$$\int_0^1 W(p)\, dp = 1.$$

Note that $V\{(R_1, \ldots, R_m)\}_p$ is

$$T_p^2 \left[\frac{p^{\frac{1}{\alpha}}(-\ln p)}{\alpha^2[-\ln(1 - p^{\frac{1}{\alpha}})](1 - p^{\frac{1}{\alpha}})}, \quad -\frac{1}{\lambda} \right] \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} \begin{bmatrix} \frac{p^{\frac{1}{\alpha}}(-\ln p)}{\alpha^2[-\ln(1 - p^{\frac{1}{\alpha}})](1 - p^{\frac{1}{\alpha}})} \\ -\frac{1}{\lambda} \end{bmatrix}.$$

It may be noted that criterion (12) is very flexible and it may take care over all variability of the percentile estimator due to a particular censoring scheme. Depending on the choice of $W(p)$, it can be made dependent or independent of p. For example, if we choose $W(p) = 1$ for all $0 < p < 1$, then the information measure $I_W\{(R_1, \ldots, R_m)\}$ is independent of p. Moreover, if we want to consider the variability of the percentile estimator for certain ranges of p, we can choose $W(p)$ accordingly. Clearly, this is a more general information measure than those information measures proposed by Zhang and Meeker (2005) or Ng et al. (2004). Using $W(p)$ as the point mass at any

Table 6 The optimal censoring scheme for different criteria when $\alpha = 2$, $\lambda = 1$, $m = 5$ and $n = 10, 15,$ 25, and 30. Against each criterion, the first row, second row, third row, and fourth row represent the optimal censoring schemes corresponding to $n = 10$, $n = 15$, $n = 25$, and $n = 30$, respectively

Criterion	R_1	R_2	R_3	R_4	R_5
1	0	0	0	0	5
	0	0	0	0	10
	0	0	0	5	15
	0	0	0	6	19
2	5	0	0	0	0
	0	0	2	8	0
	0	0	1	0	19
	0	0	0	0	25
3	5	0	0	0	0
	10	0	0	0	0
	0	0	1	0	19
	0	0	1	2	22
4	0	0	0	5	0
	0	0	0	0	10
	0	0	0	0	20
	0	0	0	6	19
5	0	0	0	0	5
	0	0	0	0	10
	0	0	0	0	20
	0	0	0	6	19
6	0	0	0	0	5
	0	0	0	0	10
	0	0	0	5	15
	0	0	0	6	19

particular point, the information measures proposed by Zhang and Meeker (2005) or Ng et al. (2004) can be obtained.

Now we provide few optimal censoring schemes for different criteria. We have used the minimum trace criterion (Criterion-1), minimum determinant criterion (Criterion-2), and the minimum variance of the pth percentile estimator for different p, namely $p = 0.5$ (Criterion-3), $p = 0.9$ (Criterion-4), and $p = 0.999$ (Criterion-5). Finally, we have used minimizing $I_W\{(R_1, \ldots, R_m)\}$ as defined in (12) (Criterion-6). Note that, in all the above cases, the minimization has to be performed numerically. They are discrete optimization problems. For given n, m, α, and λ, the optimum censoring scheme with respect to a given criterion can be found by exhaustive search for all possible R_i values satisfying (11). For illustrative purpose, the results are presented in Tables 6 and 7 when $\alpha = 2.0$ and $\lambda = 1$. Interestingly, it is

Table 7 The optimal censoring scheme for different criteria when $\alpha = 2$, $\lambda = 1$, $m = 6, 8, 10$, and $n = 15$

Table-7a: $[n = 15, m = 6]$

Criterion	R_1	R_2	R_3	R_4	R_5	R_6
1	0	0	0	0	2	7
2	0	8	1	0	0	0
3	9	0	0	0	0	0
4	0	0	0	0	1	8
5	0	0	0	0	1	8
6	0	0	0	0	2	7

Table-7b: $[n = 15, m = 8]$

Criterion	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8
1	0	0	0	0	0	0	0	7
2	6	1	0	0	0	0	0	0
3	6	1	0	0	0	0	0	0
4	0	0	0	3	2	2	0	0
5	0	0	0	0	0	0	4	3
6	0	0	0	0	0	0	0	7

Table-7c: $[n = 15, m = 10]$

Criterion	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	R_{10}
1	0	0	0	0	0	0	0	0	0	5
2	4	1	0	0	0	0	0	0	0	0
3	4	1	0	0	0	0	0	0	0	0
4	4	1	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	5	0	0
6	0	0	0	0	0	0	0	0	0	5

observed that, in all the cases considered, the Criterion-1 and Criterion-6 provide the same optimal censoring schemes.

Although the total number of sampling schemes are finite, they can be quite large. For fixed m and n, total $\binom{n-1}{m-1}$ possible progressive censoring schemes are available. For example, when $n = 25$ and $m = 12$, then the possible number of censoring schemes is $\binom{24}{11} = 2496144$, which is quite large. Till date, we do not have any efficient algorithm to find the optimal censoring scheme in this case. We propose the following sub-optimal censoring scheme. Note that, for fixed n and m, all the censoring schemes of the form $(R_1, \ldots, R_m)$ such that $R_1 + \cdots + R_m = n - m$ will belong to the convex hull generated by the points $(n - m, 0, \ldots, 0), \ldots, (0, \ldots, 0, n - m)$. Therefore, a sub-optimal censoring scheme can be obtained by choosing the optimal censoring scheme among these extreme points on the convex hull.

6 Conclusions

In this paper, we have considered the estimation of the GE parameters when the data are progressively censored. It is observed that when both the parameters are unknown, the maximum likelihood estimates cannot be obtained in explicit form. In our simulation studies, it is observed that, for certain cases, the Newton–Raphson algorithm does not converge. We have used the EM algorithm to compute the MLEs of the unknown shape and scale parameters and observed their performance through numerical simulations. It is observed that the proposed EM algorithm works quite well.

We have also proposed different criteria to compare two different sampling schemes based on their information contents. We have also reported the optimal sampling schemes with respect to different criteria for small values of m and n. It is an important problem to find an algorithm to choose the optimal sampling scheme. More work is needed in this direction.

Appendix

To prove (5), we need the following theorem.

Theorem 1 *Given $X_{(1)} = x_{(1)}, \ldots, X_{(j)} = x_{(j)}$, the conditional distribution of Z_{jk} for $k = 1, \ldots, R_j$ is*

$$f_{Z|X}(z_j|X_{(1)} = x_{(1)}, \ldots, X_{(j)} = x_{(j)}) = f_{Z|X}(z_j|X_{(j)} = x_{(j)}) = \frac{f(z_j|\alpha, \lambda)}{[1 - F(x_{(j)}|\alpha, \lambda)]}$$

for $z_j > x_{(j)}$ and 0 otherwise.

Proof The proof is straightforward. For details, see Ng et al. (2002).

Note that using Theorem 1, we can write

$$E(Z_{jk}|Z_{jk} > c) = \frac{\alpha\lambda}{1 - F(c|\alpha, \lambda)} \times \int_c^\infty x e^{-\lambda x}\left(1 - e^{-\lambda x}\right)^{\alpha-1} dx \quad \text{(put } \lambda x = y\text{)}$$

$$= \frac{\alpha}{\lambda(1 - F(c|\alpha, \lambda))} \times \int_{c\lambda}^\infty y e^{-y}\left(1 - e^{-y}\right)^{\alpha-1} dy \quad \text{(put } e^{-y} = z\text{)}$$

$$= -\frac{\alpha}{\lambda(1 - F(c|\alpha, \lambda))} \int_0^{e^{-c\lambda}} \ln z (1 - z)^{\alpha-1} dz$$

$$= -\frac{\alpha}{\lambda(1 - F(c|\alpha, \lambda))} u(c, \alpha, \lambda).$$

To prove (6), note that using Theorem 1, we can write

$$E\left(\ln\left(1 - e^{-\lambda Z_{jk}}\right)|Z_{jk} > c\right)$$

$$= \frac{\alpha\lambda}{1 - F(c|\alpha, \lambda)} \times \int_c^\infty \ln\left(1 - e^{-\lambda x}\right) e^{-\lambda x} \left(1 - e^{-\lambda x}\right)^{\alpha-1} dx$$

$$= \frac{1}{\alpha(1 - F(c|\alpha, \lambda))}$$

$$\times \int_{(1-e^{-c\lambda})^\alpha}^1 \ln y \, dy \quad \left(\text{using } y = \left(1 - e^{-\lambda x}\right)^\alpha\right)$$

$$= \frac{1}{\alpha(1 - F(c|\alpha, \lambda))}$$

$$\times \left[\left(1 - e^{-c\lambda}\right)^\alpha \left(1 - \alpha \ln\left(1 - e^{-c\lambda}\right)\right) - 1\right]. \qquad \square$$

References

Abramowitz M, Stegun IA (1964) Handbook of mathematical functions with formulas, graphs and mathematical tables. Dover, New York

Balakrishnan N (2007) Progressive censoring methodology: an appraisal. TEST 16(2):211–296 (with discussions)

Balakrishnan N, Aggarwala R (2000) Progressive censoring, theory, methods and applications. Birkhauser, Boston

Balakrishnan N, Kannan N (2001) Point and interval estimation for parameters of the logistic distribution based on progressively type-II censored data. In: Handbook of Statistics, vol 20. North Holland, Amsterdam

Balakrishnan N, Sandhu RA (1995) A simple algorithm for generating progressively type-II generated samples. Am Stat 49:229–230

Balakrishnan N, Kannan N, Lin CT, Wu SJS (2004) Inference for the extreme value distribution under progressive type-II censoring. J Stat Comput Simul 25–45

Balasooriya U, Balakrishnan N (2000) Reliability sampling plan for log-normal distribution. IEEE Trans Reliab 49:199–203

Cohen AC (1963) Progressively censored samples in life testing. Technometrics 5:327–329

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via EM algorithm. J R Stat Soc, Ser B 39:1–38

Gupta RD, Kundu D (1999) Generalized exponential distribution. Aust N Z J Stat 41:173–188

Gupta RD, Kundu D (2001) Generalized exponential distribution: different methods of estimations. J Stat Comput Simul 69:315–338

Gupta RD, Kundu D (2006) Comparison of the Fisher information matrices of the Weibull and GE distributions. J Stat Plan Inference 136(9):3130–3144

Gupta RD, Kundu D (2007) Generalized exponential distribution; existing methods and some recent developments. J Stat Plan Inference 137:3537–3547

Kundu D (2007) On hybrid censored Weibull distribution. J Stat Plan Inference 137:2127–2142

Lawless JF (1982) Statistical models and methods for lifetime data. Wiley, New York

Little RJA, Rubin DB (1983) Incomplete data. In: Kotz S, Johnson NL (eds) Encyclopedia of statistical sciences, vol 4. Wiley, New York, pp 46–53

Louis TA (1982) Finding the observed information matrix using the EM algorithm. J R Stat Soc, Ser B 44:226–233

Mann NR (1971) Best linear invariant estimation for Weibull parameters under progressive censoring. Technometrics 13:521–533

Ng T, Chan CS, Balakrishnan N (2002) Estimation of parameters from progressively censored data using EM algorithm. Comput Stat Data Anal 39:371–386

Ng T, Chan CS, Balakrishnan N (2004) Optimal progressive censoring plans for the Weibull distribution. Technometrics 46:470–481

Tanner MA (1993) Tools for statistical inferences: observed data and data augmentation methods, 2nd edn. Springer, New York

Thomas DR, Wilson WM (1972) Linear order statistic estimation for the two-parameter Weibull and extreme value distributions from type-II progressively censored samples. Technometrics 14:679–691

Viveros R, Balakrishnan N (1994) Interval estimation of parameters of life from progressively censored data. Technometrics 36:84–91

Wang BX, Yu K (2007) Optimum plan for step-stress model with progressive type-II censoring. TEST. doi:10.1007/s11749-007-0060-z

Zhang Y, Meeker WQ (2005) Bayesian life test planning for the Weibull distribution with given shape parameter. Metrika 61:237–249